SFitter: Determining supersymmetric parameters

R. Lafaye, M. Rauch, T. Plehn, D. Zerwas

SFitter: Determining supersymmetric parameters

Rémi Lafaye[1], *Michael Rauch*[2], *Tilman Plehn*[2] *and Dirk Zerwas*[3]
[1]LAPP Annecy
[2]University of Edinburgh
[3]LAL Orsay

Abstract
If supersymmetry (or similar complex models) is found at the LHC, the goal for all colliders over the coming decades will be to extract its fundamental parameters from the measurements. Dedicated state–of–the–art tools will be necessary to link a wealth of measurements to an e.g. 20–dimensional MSSM parameter space. Starting from a general log–likelihood function of this high-dimensional parameter space we show how we can find the best–fit parameter values and determine their errors. Beyond a single best–fit point we illustrate how distinct secondary minima occur in complex parameter spaces. In cases where there are flat dimensions in the likelihood we comment on the benefits and limitations of marginalizing over additional dimensions.

1 Introduction

The LHC start is now a matter of months, and high energy physicists are eager to see the first sign of a Higgs boson or any alternative to such a fundamental particle. However, fundamental scalars naturally lead to the existence of an ultraviolet completion of the the Standard Model. Such an extension of the current Standard Model, might even at the same time solve the second main mystery of high–energy physics, the existence of cold dark matter.

Supersymmetry is one appealing extension of the Standard Model and is already constrained by previous experiments such as LEP and Tevatron. The LHC era might give many hints about new–physics scenarios and it will certainly rule out large classes of extensions to the Standard Model. However, it will not give us anything like a one–to–one map between a limited number of observables and a well–defined small set of parameters.

The SFitter[1] program aim at extracting parameters compatible with the available observables, while relying on as few assumptions as possible. For example, assuming an MSUGRA scenario, the LHC might provide sufficient measurements to give good uncertainties on the fundamental parameters, with the use of well–known fitting techniques. But to test the assumption of the GUT scale unification, one needs to scan a full 20 parameter space at the TeV scale. This task requires a subtle and careful scanning of the parameter space, and is better performed with the use of techniques such as Markov chains.

2 Fitting principles

The SFitter program uses as input high–energy physics experimental data and low–energy scale constraints and compare them to theoretical predictions to compute a likelihood value. Theoretical quantities are computed to the highest order (NLO in most cases) available thanks to SUSY spectrum calculators[2, 3]. Other quantities, such as cross–sections, branching ratios and dark matter relic density can then be derived using other available programs [4, 5, 6].

The likelihood can be computed using two different scheme, depending on the final use:

– The RFit scheme as defined by Höcker et al.[7] should be used for the sake of correctness in frequentist analysis. In this case, the theoretical errors are interpreted as a lack of knowledge on a

parameter. The combined likelihood (including both experimental uncertainty σ_{exp} and theoretical error σ_{th}) is defined as:

$$-2\ln\mathcal{L} = \begin{cases} 0, & \forall\,|x_{exp}-x_{th}| < \sigma_{th} \\ \left(\frac{|x_{exp}-x_{th}|-\sigma_{th}}{\sigma_{exp}}\right)^2 & \forall\,|x_{exp}-x_{th}| \geq \sigma_{th} \end{cases} \quad (1)$$

– The standard convolution of experimental and theoretical uncertainties. This is done assuming the theoretical error has a probability density function (either flat or Gaussian) following Bayesian statistics.

The region of the parameter space in which the likelihood is to be computed highly depends on the fitting techniques used. Ideally, a scan covering the whole parameter space region of interest, is performed to find local likelihood maxima. Then, a gradient fit around each maximum allows to determine the parameter values and errors.

The parameter space region of interest definition and how the scan is performed relies on priors. There is no way around this. But hopefully if the likelihood shape is sufficiently smooth compared to the scan steps, the gradient fit will converge to the true minima. The main problem with a 20–dimensional parameter space (like with the phenomenological MSSM) is to perform an efficient scan. Efficient in terms of coverage and computing time. This is were the Markov chains come into play.

SFitter provides all relevant frequentist or Bayesian answers in three steps: first (1), we compute a likelihood map of the entire parameter space, using either a simple grid method or a Markov Chains approach (described later). This map is completely exclusive, i.e., it includes all dimensions in the parameter space. Then (2), we rank the best local likelihood maxima in the map according to their log–likelihood values. This way we identify the global maximum, and everybody can include their personal prior towards secondary maxima (i.e., SUSY breaking scenario), without mistaking such a prior for actual likelihood. Last (3), we compute likelihood or probability maps of lower dimensionality, down to one–dimensional distributions, by properly removing or marginalizing unwanted parameter dimensions.

3 Markov Chains

A Markovian process is defined as a stochastic process in which the conditional probability distribution of future states depends only on the present state and not on any past state.

For the purpose of fitting the state is defined as a point in the parameter space and its associated likelihood value. The future parameter values (*new*) are then chosen according to the current position (*cur*) and kept (as the next point of the chain) if it satisfies one of the two following conditions:

$$\begin{cases} \mathcal{L}_{new} > \mathcal{L}_{cur} \\ \texttt{Random}[0,1] < \frac{\ln\mathcal{L}_{cur}}{\ln\mathcal{L}_{new}} \end{cases} \quad (2)$$

If the new point is chosen randomly over the whole parameter space without any dependence on the current point, then the Markov chain is equivalent to a Monte–Carlo fitting method. The main drawback in this case is that high likelihood regions will not be favored with respect to low likelihood ones unless we a–priori know the likelihood shape and can generate new points accordingly (as in Monte–Carlo resonant process generation). Another way to improve the Markov chains efficiency is to generate the new point depending on the current point. In SFitter, this is done using a Landau distribution separatly for each parameter. The Landau peak is taken as the current value of the parameter and the distribution extends to the parameter limits.

The main advantage of the Markov chains method compared to a crude scan is its convergence speed, which can go linearly with the number of parameters. Indeed, parameters which have no influence on the likelihood value do not slow down the convergence process. Also, it does not rely on the likelihood

χ^2	m_0	$m_{1/2}$	$\tan\beta$	A_0	μ	m_t
0.09	102.0	254.0	11.5	-95.2	$+$	172.4
1.50	104.8	242.1	12.9	-174.4	$-$	172.3
73.2	108.1	266.4	14.6	742.4	$+$	173.7
139.5	112.1	261.0	18.0	632.6	$-$	173.0
...						

Table 1: SFitter output for MSUGRA in SPS1a. List of the best log–likelihood values over the MSUGRA parameter space. All masses are given in GeV.

shape in the parameter space, and like any other scanning process, it has the ability to find secondary minima. However, it should be used cautiously as it is not meant to find the exact value of the minima, and a bad choice of priors can lead to scans on limited parameter space regions. Theoretically increasing the number of points in the chain can overcome these problems. Alternatively, one can try different priors to make sure the whole parameter space is correctly scanned and then use a gradient fit to find the exact minima.

4 MSUGRA as a toy model

Clearly, in the LHC era no model for supersymmetry breaking should be assumed for analysis. Instead, the breaking mechanism should be inferred from data. At the LHC (and certainly in combination with the ILC) there is little need for top–down analysis, which are known to reveal more about their author's imagination than about physics. As a matter of fact, supersymmetry should only be considered one possible interpretation of for example cascade decays. However, completely generalizing an intelligent analysis to a general new–physics model space seems not viable at the moment. We will therefore assume that supersymmetry, little–Higgs models or extra–dimensional models can be distinguished by simpler hypothesis testing.

Running the Markov chains algorithm on the MSUGRA SPS1a point[9] using 300 fb$^-1$ LHC toy data[8], we obtains the results summarized in table1. As discussed before this ranked list of likelihood maxima has to be refined with a gradient fit Building a profile likelihood map (i.e., looking for the maximum likelihood in all directions but the ones of interest) can be compared to the Bayesian approach (marginalizing over all other dimensions to obtain a probability density function). The two illustrations shown in figure 1 look similar. However, there are two differences in the details: first, the area around the true parameter point is less pronounced in the Bayesian pdf, compared to the profile likelihood. When we integrate over a direction in parameter space we largely collect noise from regions with small likelihood. This noise washes out the peaked structures. The second effect is the more pronounced branch structure for the Bayesian pdf, while in the profile likelihood the area between the two branches is filled by single good parameter points in the parameter projected away; the marginalization provides us with 'typical' likelihood values in this region which in general does not fit the data well.

The washing–out effect also considerably smears the one–dimensional Bayesian pdf distribution in m_0 on figure 2. But here the marginalization over $m_{1/2}$ also creates a higher peak at $m_0 = 50$ GeV, which should only be interpreted as higher likelihood density. The only reliable source of information on likelihood value being the profile map.

5 Conclusions

It will be hard for the LHC to give a conclusive answer to the crucial question, namely what is the 'correct' ultraviolet completion of the Standard Model. The impact of the LHC on the vast model space will at best be locally conclusive. Be it from a frequentist or a Bayesian point of view, the use of improved fitting techniques, such as used in SFitter, will prove to be very useful. At least until we reach sufficient

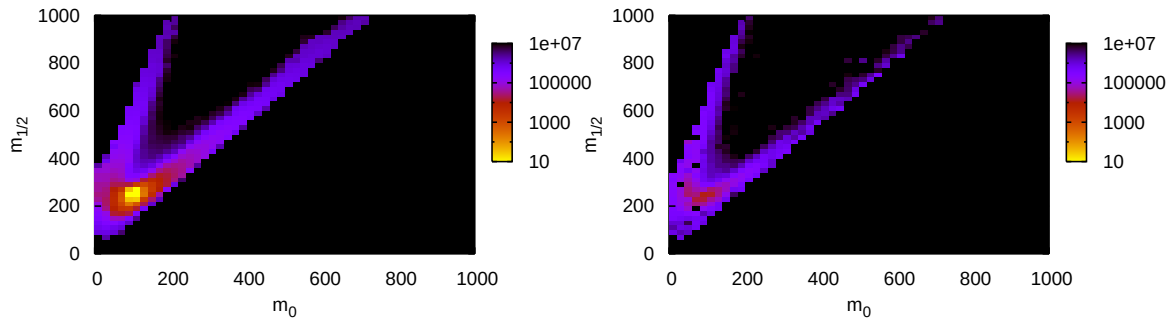

Fig. 1: SFitter output for MSUGRA in SPS1a. Left: two–dimensional profile likelihood χ^2 over the m_0–$m_{1/2}$ plane. Right: two–dimensional Bayesian pdf χ^2 over the m_0–$m_{1/2}$ plane marginalized over all other parameters. All masses are given in GeV.

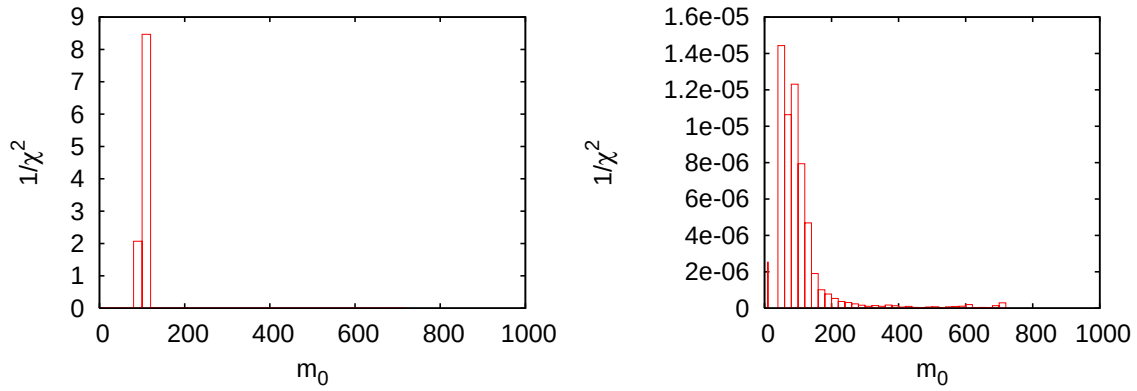

Fig. 2: SFitter output for MSUGRA in SPS1a. Left: one–dimensional profile likelihood $1/\chi^2$ for m_0. Right: one–dimensional Bayesian pdfs $1/\chi^2$ for m_0. All masses are given in GeV.

experimental and theoretical precision to narrow down the *Next Standard Model* parameters.

References

[1] R. Lafaye, T. Plehn, M. Rauch and D. Zerwas, arXiv:0709.3985 [hep-ph].

[2] A. Djouadi, J. L. Kneur and G. Moultaka, arXiv:hep-ph/0211331.

[3] B. C. Allanach, Comput. Phys. Commun. **143**, 305 (2002).

[4] W. Beenakker, R. Höpker, M. Krämer, T. Plehn, M. Spira, P. M. Zerwas, Nucl. Phys. B **492**, 51 (1997); Nucl. Phys. B **515**, 3 (1998); Phys. Rev. Lett. **83**, 3780 (1999); arXiv:hep-ph/9809319.

[5] A. Djouadi, M.M. Muhlleitner and M. Spira CERN-PH-TH/2006-200

[6] G. Bélanger, F. Boudjema, A. Pukhov and A. Semenov Comput.Phys.Commun.176:367-382,2007

[7] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, L. Roos Eur. Phys. J. C41, 1-131 (2005)

[8] G. Weiglein et al. [LHC/LC Study Group], arXiv:hep-ph/0410364.

[9] B. C. Allanach et al., in Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001) ed. N. Graf, Eur. Phys. J. C 25 (2002) 113